

Method for producing an active heterodimeric AMV-RT in prokaryotic cells

The invention concerns a method for producing a recombinant active heterodimeric AMV-RT by expressing one or several DNA sequences coding for the α - and/or β -subunit(s) of the AMV-RT in prokaryotic cells under certain growth and induction conditions.

The discovery of the reverse transcriptases in the seventies disproved the "central dogma" of molecular biology on the information transfer from DNA via RNA to protein as a unidirectional process (Termin H. and Mizutani S., 1970 Nature 226:1211-1213; Baltimore D., 1970, Nature 226:1209-1211). The enzymatic characterization of these RNA-dependent DNA polymerases is the basis for current understanding on the amplification cycle of RNA viruses and thus also on the development and spread of diseases that are caused by this type of virus (cancer, AIDS etc.).

However, reverse transcriptases are also a tool for molecular biologists for the synthesis, amplification and cloning of cDNAs (RT-PCR). This technology allows a simplified and accelerated examination of gene expression in eukaryotic cells. After isolating the total mRNA from cell extracts or tissues, the mRNA is translated back into cDNA by the reverse transcriptase and amplified by the subsequent PCR step to enable cloning and characterization. Consequently it is not necessary to, on the one hand, elucidate the intron and

exon structures of the genes but, on the other hand, it is also possible to examine gene expression in the cell during various life cycles or during the development of diseases (such as cancer).

Reverse transcriptases (RT) from three different retroviruses have hitherto been closely examined: The RT from *Moloney Murine Leukemia Virus* (M-MLV). This enzyme consists of a single subunit with a molecular weight of 78 kDa (Prasad V.R., 1993 reviewed in *Reverse Transcriptase*, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press, 135). In addition an RT from *Human Immunodeficiency Virus* (HIV) is known. This RT is a heterodimer that is composed of two subunits p66 and p51, the p51 subunit being formed by proteolytic cleavage of p66 (Le Grice S.F.J., 1993 reviewed in *Reverse Transcriptase*, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press, 163). In addition RTs are known from *Avian Sarcoma-Leukosis Virus* (ASLV). The RT obtainable from *Avian Myeloblastosis Virus* (AMV) also belongs to the ASLV family. This RT is also a heterodimer that is composed of an α -chain with a molecular weight of ca. 63 kDa and a β -chain with a molecular weight of ca. 95 kDa. In this case the α -chain is also formed by proteolytic processing of the β -chain (Golomb M. and Grandgenett D., 1979, *J. Biol. Chem.* **254**: 1606-1613; Weiss R. et al., eds. 1984, *Molecular Biology of tumor viruses*, 2nd edition: RNA tumor viruses 1/text. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York).

Whereas the M-MLV-RT is expressed in *E. coli* as a monomer and the HIV-RT as a heterodimer, it has so far not been possible to express the AMV-RT as an active or soluble heterodimer in *E. coli* or other prokaryotes to a

satisfactory degree. Although according to WO 00/42199 certain RT variants are expressed in *E. coli* or preferably in eukaryotic insect cells, the desired RT that is obtained in this process mainly consists (ca. 90 %) of an insoluble component.

In addition it is difficult to measure a recombinant AMV-RT in crude cell extracts of *E. coli* since, on the one hand, RNA templates are degraded by intrinsic *E. coli* RNases and, on the other hand, *E. coli* strains have a DNA polymerase which also has an RT activity in addition to the DNA polymerase activity (Ricchetti, M. and Huc, H., 1993, *EMBO J.* **12** (2), 387-396). Hence this intrinsic *E. coli* RT activity considerably interferes with the determination of the activity of the recombinant AMV-RT in crude *E. coli* extracts and in fractions from the purification.

Hence the object of the present invention is to provide a recombinant active heterodimeric AMV-RT in adequate amounts.

The object is achieved by a method for producing an active heterodimeric AMV-RT in prokaryotic host cells wherein one or several DNA sequence(s) which code for the α and β subunit or chain of the AMV-RT, are cloned into expression plasmids, the expression plasmids are transformed in prokaryotic cells, the expression of the heterodimeric AMV-RT is induced and the recombinant heterodimeric AMV-RT is purified i.e. isolated from the cells. Suitable genes and DNA sequences are, among others, those which only code for one of the AMV-RT subunits. A portion or the expression product can subsequently be converted by certain measures, such as

proteolytic cleavage of the β -chain, into the other subunit. The sequences SEQ ID NO:4 and SEQ ID NO:5 have proven to be particularly suitable for the method according to the invention which generates an active heterodimeric AMV-RT composed of the subunits SEQ ID NO:6 and SEQ ID NO:7.

The structural genes and DNA sequences coding for the subunits of the AMV-RT can either be cloned on different, separate expression plasmids or on one expression plasmid, optionally in the presence of so-called helper plasmids, and expressed in a suitable host cell. Suitable expression plasmids are for example pDS, pKK177-3 or pKKT5. The plasmid pKKT5 in which the respective structural genes are inserted under the control of the T5 promoter is preferred according to the invention. Other potential promoters, which are preferably IPTG-inducible promoters, are for example the lac, lac UV5 or tac promoter. Alternative helper plasmids such as the plasmid pUBS520 and suitable selection markers such as ampicillin or kanamycin are in principle known to a person skilled in the art.

The expression plasmids and optionally other helper plasmids are transformed into a suitable prokaryotic host cell. According to the invention it is preferable to use an *E. coli* strain such as *E. coli* K12 C600, DH5 α , LE392, JM83, JM105, NM522, M15, RR1 Δ 15, UT5600, TG1, A1200 or the strains *E. coli* B, BL21, HB101. The *E. coli* strain LE392 is particularly preferred according to the invention.

The expression of the heterodimeric AMV-RT can be induced by various measures. In particular certain

growth and induction conditions have positive effects on the expression of active AMV-RT. A growth temperature in the range of 10° to 25°C combined with a low inducer concentration has proven to be advantageous according to the invention. A growth temperature of about 15°C and an inducer concentration between 0.1 and 0.5 mM, preferably of about 0.15 mM, have proven to be particularly suitable. IPTG (isopropyl- β -D-thiogalactopyranoside) or lactose are preferably used according to the invention as the inducer.

Furthermore it turned out that the soluble expression of AMV-RT in prokaryotic cells can be increased by the co-expression of helper genes. Potential helper genes are in particular the trpT gene which codes for the tryptophan tRNA. In addition chaperone genes are suitable for soluble expression such as the genes coding for GroEL and GroES, GrpE, ClpB, Dnak and DnaJ. The genes for one or several chaperones are then preferably located on a helper plasmid with an inducible promoter; the genes which code for the chaperones GroEL and GroES are under the control of a constitutive promoter on the expression plasmid on which the structural genes for the α and/or β chain are also located. However, it is particularly preferred according to the invention when the genes coding for GroEL and GroES are cloned on the expression plasmid which carries the genes for the α - and β -chain and the genes coding for Dnak, DnaJ, GrpE and ClpB are cloned on a helper plasmid.

In addition to methods that are generally known to a person skilled in the art, it is especially advantageous to use affinity chromatography materials such as metal ion chelating materials or cation exchangers to purify and isolate the recombinant heterodimeric AMV-RT from

the cell extract. It is particularly advantageous for the purification of the AMV-RT for the expression products, i.e. the α - as well as the β -chain to be fused with peptide sequences that are able to reversibly bind to particular column materials such as cation exchangers, metal ion chelating materials such as nickel, copper or zinc nitriloacetic acid (NTA) resins. Peptide sequences that are suitable according to the invention can have from two to about 100 amino acids or amino acid derivatives. Peptide sequences which are composed of two to ten amino acids, e.g. arginine residues or histidine residues, have proven to be particularly suitable for the invention. In addition it has also proven to be particularly advantageous to use such peptide sequences comprising eight arginine or six histidine residues. In addition commercially available peptide sequences such as Strep-tag[®] (IBA GmbH, Göttingen/Germany) or GST-tag (Pharmacia, Uppsala/Sweden) are also suitable for the method according to the invention.

The invention is further elucidated by the following examples:

1. Example:

Isolation of genes which code for the α - chain and β -chain

The data bank sequence (MEDLINE ID 94366722, Baluda *et al.*, 1994) was used to design oligonucleotide primers for the isolation of the β -chain (see SEQ ID NO:1 and 2). A *Eco*RI restriction endonuclease cleavage site was incorporated at the 5' end and a *Pst*I restriction cleavage site was incorporated at the 3' end for the

subsequent cloning into vectors. In addition a further 3' primer was designed (see SEQ ID NO:3) which enables the isolation of the α -chain. Both chains were fished by means of PCR from a virus lysate (ATCC VR-265) by means of RT-PCR as well as from an *E. coli* clone (ATCC 31990) which carries the β -chain on a plasmid. The PCR mixtures were applied to a 1 % agarose gel, the PCR fragments of ca. 1715 bp for the α -chain and ca. 2570 bp for the β -chain were isolated from the agarose gel (QIAEX II, Gel Extraction Kit, Qiagen/Germany), cleaved with the restriction endonucleases mentioned above and cloned into a vector fragment of pUC19 that had also been linearized with *Eco*RI and *Pst*I and isolated. For this 1 μ l (20 ng) vector fragment and 3 μ l (100 ng) PCR fragment, 1 μ l 10 x ligase buffer (Maniatis et al., 1989 Molecular Cloning: A Laboratory Manual, second Edition, Cold Spring Harbor Laboratory Press NY (USA), vol. 27), 1 μ l T4 DNA ligase, 4 μ l sterile H₂O bidistilled were pipetted, carefully mixed and incubated overnight at 16°C. The cloned genes were subsequently examined by means of restriction analysis and sequencing. The sequences are shown in SEQ ID NO:4 (α -chain) and SEQ ID NO: 5 (β -chain).

Comparison with the data bank sequence (MEDLINE ID 94366722, Baluda, M.A., and Reddy, E.P., 1994, Oncogene 9:2761-2774) yielded a homology of 98.8 % at the DNA level for the α -chain as well as for the β -chain. When the resulting amino acid sequences are compared, it becomes apparent that most substitutions at the DNA level are so-called silent mutations i.e. do not lead to amino acid substitutions. Only three base substitutions also resulted in amino acid substitutions but they are found reproducibly in each isolated PCR product. These are the substitutions Arg273Met, Arg304Gln and

Asp495Glu. The amino acid sequences of both chains are shown in SEQ ID NO:6 (α -chain) and SEQ ID NO:7 (β -chain).

2. Example:

Expression of the α -chain and β -chain without fused peptide sequences (tags)

2.1. Construction of the expression plasmids pAMV- α and pAMV- β

In order to express the AMV-RT, the genes for both chains were cloned separately into expression vectors in such a manner that the structural genes were each inserted in the correct orientation under the control of the T5 promoter. For this the respective structural gene for the α -chain and the β -chain were cut out of the plasmid pUC19 by EcoRI and PstI, the restriction mixtures were separated by agarose gel electrophoresis and the 1715 bp fragment of the α -chain and the 2570 bp fragment of the β -chain were isolated from the agarose gel. The expression plasmid pKKT5, which is formed from pKK177-3 (Kopetzki et al., 1989, *Mol. Gen. Genet.* **216**: 149-155) by replacing the tac promoter with the T5 promoter from pDS (Bujard et al., 1987, *Methods Enzymol.* **155**: 416-433), was used for the expression. The EcoRI restriction endonuclease cleavage site in the sequence of the T5 promoter was removed by two point mutations. The resulting expression plasmid was cut with EcoRI and PstI for the insertion of the genes for the AMV-RT, the restriction mixture was separated by agarose gel electrophoresis and the resulting vector fragment of ca. 2500 bp was isolated from the agarose gel. The vector fragment obtained in this manner was separately ligated

as described above with the genes for the α -chain and the β -chain described in example 1. The correct insertion of the genes was checked by restriction control and sequencing. The resulting plasmid pAMV- α and pAMV- β was firstly separately cotransformed with the helper plasmid pUBS520 for expression control in various *E. coli* strains. It is conceivable in this case that the α -chain and the β -chain could be separately expressed in order to obtain $\alpha\alpha$ - and $\beta\beta$ -homodimers. The helper plasmid pUBS520 (Brinkmann et al., 1989, *Gene* **85**: 109-114) carries inter alia the *lacI^Q* gene which codes for the lac repressor and the *dnaY* gene which codes for the rare tRNA^{Arg} in *E. coli* which recognizes the codons AGA and AGG (Garcia et al., 1986, *Cell* **45**: 453-459). The kanamycin resistance gene from the transposon TN903 was used as a selection marker.

2.2 Separate transformation of the expression plasmids pAMV- α and pAMV- β in *E. coli*

Competent cells of various *E. coli* strains were prepared according to the method of Hanahan (*J. Mol. Biol.* 1983, vol. **166**, 557). 200 μ l of *E. coli* LE392 cells prepared in this manner were admixed with 20 ng isolated expression plasmid pAMV- α DNA or pAMV- β DNA and 40 ng helper plasmid DNA. After 30 min incubation on ice a heat shock (90 sec. at 42°C) was carried out. Subsequently the cells were transferred to 1 ml LB medium and incubated for 1 hour at 37°C in LB medium for the phenotypic expression. Aliquots of this transformation mixture were plated out on LB plates containing ampicillin and kanamycin as selection markers and incubated for 15 hours at 37°C.

2.3 Expression of the gene for the α -chain in *E. coli*

In order to express the gene which codes for the α -chain of the AMV-RT, plasmid-containing clones were inoculated in 3 ml LBampkan medium and incubated at 30°C in a shaker. At an optical density of 0.5 (measured at 550 nm, OD₅₅₀) the cells were induced with 0.5 mM IPTG and incubated for 4 h at 30°C in a shaker. Subsequently the optical density of the individual expression clones was determined, an aliquot which corresponded to an OD₅₅₀ of 5.0/ml was removed and the cells were centrifuged (10 min, 6000 rpm, 4°C). The cell pellet was resuspended in 400 μ l TE buffer (50 mM TRIS/50 mM EDTA, pH 8.0), the cells were disrupted by ultrasound and the soluble protein fraction was separated from the insoluble protein fraction by centrifugation (10 min, 14000 rpm, 4°C). Application buffer containing SDS and β -mercaptoethanol was added to all fractions and the proteins were denatured by boiling (5 min 100°C). Afterwards 10 μ l of each was analysed by means of an analytical SDS gel (10 %) (Laemmli U.K., 1970, *Nature* 227: 555-557).

Analysis of the SDS gel shows a clear overexpression of the α -chain. A strongly overexpressed additional band is seen at ca. 63 kDa which is not observed with the non-induced control clones or the induced control clones which do not contain plasmid. A small portion of the overexpressed α -chain appears in the soluble protein fraction whereas the major amount is formed as an insoluble expressed protein.

2.4 Expression of the gene of the β -chain in *E. coli*

In order to express the gene which codes for the β -chain

of the AMV-RT, 3 ml LBampkan medium was inoculated with plasmid-containing clones and incubated at 30°C in a shaker. At an OD_{550nm} of 0.5 the cells were induced with 0.5 mM IPTG and incubated for 4 h at 30°C in a shaker. Subsequently the optical density of the individual expression clones was determined, an aliquot which corresponded to an OD₅₅₀ of 5.0/ml was removed and the cells were centrifuged (10 min, 6000 rpm, 4°C). The cell pellet was resuspended in 400 µl TE buffer (50 mM TRIS/50 mM EDTA, pH 8.0), the cells were disrupted by ultrasound and the soluble protein fraction was separated from the insoluble protein fraction by centrifugation (10 min, 14000 rpm, 4°C). Application buffer containing SDS and β-mercaptoethanol was added to all fractions and the proteins were denatured by boiling (5 min 100°C). Afterwards 10 µl of each was analysed by means of an analytical SDS gel (8 %) (Laemmli U.K., 1970, *Nature* 227: 555-557).

Analysis of the SDS gel shows a clear overexpression of the β-chain. A strongly overexpressed additional band is seen at ca. 95 kDa which is not observed with the non-induced control clones or the induced control clones which do not contain plasmid. The majority of the overexpressed β-chain appears in the insoluble protein fraction, however, a slight overexpression is also seen in the soluble protein fraction.

2.5 Expression of both chains on separate plasmids in a cell

In order to express both chains in one cell, the *lacIq* expression cassette and the *dnaY* expression cassette must firstly be recloned from the helper plasmid pUBS520

onto the expression plasmids. The *lacI^q* expression cassette was cloned onto pAMV- α and the *dnaY* expression cassette was cloned onto the expression plasmid pAMV- β . In order to ensure a stable multiplication of the expression plasmids, the ampicillin resistance gene from pAMV- α was replaced by the kanamycin resistance gene from pUBS520. The resulting expression plasmids pAMV- α *lacI^q* and pAMV- β _{*dnaY*} were subsequently cotransformed in various *E. coli* expression strains.

In order to express the genes which code for the α -chain and the β -chain of the AMV-RT, 3 ml LBampkan medium was inoculated with plasmid-containing clones and incubated at 30°C in a shaker. At an OD_{550nm} of 0.5 the cells were induced with 0.5 mM IPTG and incubated for 4 h at 30°C in a shaker. Subsequently the optical density of the individual expression clones was determined, an aliquot which corresponded to an OD₅₅₀ of 5.0/ml was removed and the cells were centrifuged (10 min, 6000 rpm, 4°C). The cell pellet was resuspended in 400 μ l TE buffer (50 mM TRIS/50 mM EDTA, pH 8.0), the cells were disrupted by ultrasound and the soluble protein fraction was separated from the insoluble protein fraction by centrifugation (10 min, 14000 rpm, 4°C). Application buffer containing SDS and β -mercaptoethanol was added to all fractions and the proteins were denatured by boiling (5 min 100°C). Afterwards 10 μ l of each was analysed by means of an analytical SDS gel (8 %) (Laemmli U.K., 1970, *Nature* 227: 555-557).

Analysis of the SDS gel surprisingly shows a clear overexpression of the α - and β -chain. Strongly overexpressed additional bands are seen at ca. 63 kDa and ca. 95 kDa which are not observed with the non-induced control clones or the induced control clones which do not

contain plasmid. The distribution of the bands in the soluble and insoluble fraction is like that of the experiments in which both chains were expressed separately. The expression output of both chains is overall somewhat less than for separate expression.

3. Example:

Expression of the α -chain and β -chain with fused tags to simplify the purification

3.1 Production of various fusion proteins

In order to efficiently purify the recombinant AMV-RT heterodimers, suitable peptide sequences, so-called tags were fused to the 5' end of both chains. Tags enable affinity chromatographies to be carried out. A series of two affinity chromatographies which are each specific for one of the two tags additionally allows the isolation of pure heterodimers (Wende W. et al., 1996, *Biol. Chem.* 377, 625-632). Appropriate primer designs were used to attach eight arginine residues to the α -chain and six histidine residues to the β -chain by means of PCR reactions. The sequences of the sense primers are shown in SEQ ID NO:8 (5' primer for the α -chain) and SEQ ID NO:9 (5' primer for the β -chain). The oligonucleotides of SEQ ID NO:2 (β -chain) and SEQ ID NO:3 (α -chain) which had already been used for gene isolation were used as antisense primers.

The PCR mixtures were applied to a 1 % agarose gel, the PCR fragments of 1739 bp for the α -chain and 2597 bp for the β -chain were isolated from the agarose gel (QIAEX II, Gel Extraction Kit, Qiagen, Germany), cleaved with the restriction endonucleases EcoRI and PstI and cloned

into a vector fragment of the preferred expression plasmid that had also been linearized with EcoRI and *Pst*I and isolated. For this 1 μ l (20 ng) vector fragment and 3 μ l (100 ng) PCR fragment, 1 μ l 10 x ligase buffer (Maniatis et al., 1989 Molecular Cloning: A Laboratory Manual, second Edition, Cold Spring Harbor Laboratory Press NY (USA), vol. 27), 1 μ l T4 DNA ligase, 4 μ l sterile H₂O_bidistilled were pipetted, carefully mixed and incubated overnight at 16°C. The cloned genes were subsequently examined by means of restriction analysis and sequencing. The resulting expression plasmids were named pAMV- α_{lacIq} -Arg and pAMV- β_{dnay} -His.

3.2 Transformation of the expression plasmids pAMV- α_{lacIq} -Arg and pAMV- β_{dnay} -His in various *E. coli* expression strains

Competent cells of various *E. coli* strains were prepared according to the method of Hanahan (*J. Mol. Biol.* 1983, vol. 166 pp. 557) (see example 2.2).

3.3 Expression of both chains with fused tags on separate plasmids in a cell

In order to express both chains with tags in a cell, various *E. coli* expression strains were cotransformed with the expression plasmids pAMV- α_{lacIq} -Arg and pAMV- β_{dnay} -His.

In order to express the genes which code for the α -chain with an Arg-tag and the β -chain with an His-tag of the AMV-RT, 3 ml LBampkan medium was inoculated with plasmid-containing clones and incubated at 30°C in a shaker. At an OD₅₅₀ of 0.5 the cells were induced with 0.5 mM IPTG and incubated for 4 h at 30°C in a shaker. Subsequently

the optical density of the individual expression clones was determined, an aliquot which corresponded to an OD₅₅₀ of 5/ml was removed and the cells were centrifuged (10 min, 6000 rpm, 4°C). The cell pellet was resuspended in 400 µl TE buffer (50 mM TRIS/50 mM EDTA, pH 8.0), the cells were disrupted by ultrasound and the soluble protein fraction was separated from the insoluble protein fraction by centrifugation (10 min, 14000 rpm, 4°C). Application buffer containing SDS and β-mercaptoethanol was added to all fractions and the proteins were denatured by boiling (5 min 100°C). Afterwards 10 µl of each was analysed by means of an 8 % analytical SDS gel (Laemmli U.K., 1970, *Nature* **227**: 555-557).

Analysis of the SDS gel surprisingly shows a clear overexpression of the α- and β-chain. Strongly overexpressed additional bands are seen at ca. 63 kDa and ca. 95 kDa which are not observed with the non-induced control clones or the induced control clones which do not contain plasmid. The distribution of the bands in the soluble and insoluble fraction is like that of the experiments in which both chains were expressed separately without tags in one cell.

3.4 Expression of both chains with fused tags on a plasmid

If the genes for the α- and β-chain of the AMV-RT are distributed on two plasmids, differences in the stability of these plasmids could lead to the production of different amounts of the respective chains and thus to a lower yield of αβ-chain heterodimer. Hence with the exception of the gene for β lactamase, the entire genetic information of the two plasmids pAMV-α_{lacIq}-Arg and pAMV-β_{dNAY}-His was combined on a single plasmid pAMV-

$\alpha\beta$ -1. This plasmid was constructed by inserting the *SspI-AfI*III fragment of pAMV- β_{dnAY} -His containing the sequence for the T5 promoter, the gene coding for the β -chain with an N-terminal His tag, the sequence for the *rrnB* terminator and the *dnaY* gene, into the SalI cleavage site of pAMV- α_{lacIq} -Arg which contains the sequence for the T5 promoter, the gene coding for the α -chain with an N-terminal Arg-tag, the sequence for the *rrnB* terminator, the kanamycin resistance gene and the *lacIq* gene. For this purpose 1 μ g each of the expression plasmids pAMV- α_{lacIq} -Arg and pAMV- β_{dnAY} -His were cleaved with the restriction endonucleases described above according to the manufacturer's instructions, the restriction mixtures were separated in a 1 % agarose gel and the 4124 bp *SspI-AfI*III fragment of pAMV- β_{dnAY} -His and the 6024 bp fragment of pAMV- α_{lacIq} -Arg were isolated from the agarose gel (QIAEX II, Gel Extraction Kit, Qiagen/Germany). The non-compatible ends were prepared with Klenow polymerase (Roche Diagnostics GmbH) according to the manufacturer's instructions and the two fragments were ligated together as described above. The resulting new expression plasmid pAMV $\alpha\beta$ -1 was examined by means of restriction analysis.

The correct expression plasmid according to restriction analysis was transformed in the *E. coli* K-12 strain LE392 as described above and subjected to an expression control. The protein content of the cells after 4 hours growth under induced conditions was subsequently examined by means of SDS-PAGE. According to SDS-PAGE analysis the level of the expression output and the relative proportion of soluble and insoluble fractions are comparable to the expression of the genes for the α - and β -chain on separate plasmids, but the amount of expressed α - and β -chain appears to be more homogeneous.

Furthermore the Arg-tag of the α -chain was replaced by a His-tag like that of the β -chain for the purification procedure. For this purpose an intermediate construct pAMV- α_{lacIq} -His was prepared in which the EcoRI-NheI fragment from pAMV- α_{lacIq} -Arg was replaced by the EcoRI-NheI fragment from pAMV- β_{dnay} -His. Subsequently, like the construction of pAMV $\alpha\beta$ -1, the entire genetic information of the two plasmids pAMV- α_{lacIq} -His and pAMV- β_{dnay} -His with the exception of the gene for β lactamase was combined on a single plasmid pAMV $\alpha\beta$ -2. The new expression vector was named pAMV $\alpha\beta$ -2. Cells were transformed as described above with pAMV $\alpha\beta$ -2 and subjected to an expression control under standard conditions. The expression output was not increased under these conditions.

4. Example:

Expression optimization

4.1 Increase of the expression of active AMV-RT by changing the expression conditions

Particular growth and induction conditions have positive effects on the expression of active AMV-RT. Afterwards the growth temperature was lowered from 30°C to 15°C during the induction phase, the IPTG concentration was reduced from 0.5 mM to 0.15 mM to induce expression and the induction time was increased from 4 h to 26 h. The protein content of the cells after the induction phase was examined as described above by SDS polyacrylamide gel electrophoresis.

Afterwards the total expression yield of α - and β -chain was, as expected, substantially reduced in the SDS-PAGE analysis, but the content of soluble expressed α - and β -

chain was considerably increased in comparison to the expression experiments under standard growth and induction conditions. This increase in the expression of active AMV-RT was also confirmed in the subsequent purification and activity determination.

4.2 Increasing the expression of active AMV-RT by coexpression of helper genes

4.2.1. Coexpression of the gene for the tryptophan-tRNA (tRNA^{trp})

One property of the AMV-RT is to use an endogenous cell tRNA for tryptophan (tRNA^{trp}) as a primer for the polymerase reaction after infection of a eukaryotic host cell (Leis et al., 1993, in: Reverse Transcriptase, Cold Spring Harbor Monograph Series, eds.: Skala, A.M. and Goff, S.P., Cold Spring Harbor NY (USA), 33-48). However, whether the endogenous *E. coli* tRNA^{trp} can be used by the AMV-RT as a primer has not been proven. In *E. coli* the tRNA^{trp} is only coded by a single gene *trpT*, the expression of which is adapted to the normal requirements of the cell. In order to exclude a potential deficiency of tRNA^{trp} in the cell, the *trpT* gene according to SEQ ID NO:10 was isolated by means of PCR from *E. coli* LE392 (the primers used for this are shown in SEQ ID NO:11 and 12), recleaved with *PstI* for insertion into pAMV- α_{lacIq} -His and ligated into the vector fragment of pAMV- α_{lacIq} -His that was also linearized with *PstI* as described above. Clones which have integrated the *trpT* gene at the *PstI* restriction endonuclease cleavage site were checked by means of restriction analysis and sequencing. In this intermediate construct pAMV- α_{lacIq} -His-*trpT* the gene for

the α -chain and the gene for the *E. coli* tRNA^{trp} form one transcription unit, the expression of which is regulated by the IPTG-inducible T5 promoter.

Subsequently, similarly to the construction of pAMV $\alpha\beta$ -1 or pAMV $\alpha\beta$ -2, the entire genetic information of the two plasmids pAMV- α_{lacIq} -His-trpT and pAMV- β_{dnay} -His with the exception of the gene for β -lactamase was combined on a single plasmid pAMV $\alpha\beta$ -3. Cells were transformed as described above with pAMV $\alpha\beta$ -3 and subjected to an expression control using the modified expression conditions. Afterwards the yield of active AMV-RT is significantly increased.

4.2.2. Coexpression of chaperone genes

In *E. coli* there are two main chaperone systems comprising the GroESL machinery and a 4 component system consisting of DnaK, DnaJ, GrpE and ClpB (Kedzierska, 1999). Both systems play an important role in the correct folding of newly formed proteins as well as in the renaturing of proteins that have aggregated as a result of stress (Hartl F.U., 1996, Nature 381, 571-580; Bukau H. and Horwitz A.L., 1998, Cell 92, 351-366; MogK A. et al., EMBO J. 18, 6934-6949; Zolkiewski M., 1999, J. Biol. Chem. 274, 28083-28086; Goloubinoff P. et al., 1999, Proc. Natl. Acad. Sci. USA 96, 13732-13737).

In a first step the groESL operon from *E. coli* should be overexpressed in the AMV-RT production strains. For this the EcoRI-HindIII fragment from pOF39 (Fayet O., Louarn J.-M., Georgopoulos C., 1986, Mol. Gen. Genet. vol. 202, pp. 335-345 was integrated in the SspI cleavage site of the plasmid pAMV- β_{dnay} -His. Non-compatible ends were prepared with Klenow polymerase (Roche Diagnostics)

according to the manufacturer's instructions before ligation. The sequence of *groESL* is shown in SEQ ID NO:13. In this new construct pAMV- β_{dnay} -His-groESL, the *groESL* operon forms an artificial transcription unit containing the 3'-situated gene for β lactamase. The expression is then either under the control of the endogenous *bla* constitutive promoter that is now on the 5' side of the *groESL* operon and/or under the control of the σ^{32} -dependent promoter of the *groESL* operon. Subsequently the entire genetic information of the two expression plasmids pAMV- α_{lacIq} -His-*trpT* and pAMV- β_{dnay} -His-groESL with the exception of the gene for β -lactamase was again combined as described above on a single plasmid pAMV $\alpha\beta$ -4.

Cells were transformed with pAMV $\alpha\beta$ -4 as described above and subjected to an expression control under the modified expression conditions. The co-overproduction of GroESL results in an increase of the biomass and of the amount of active AMV-RT. Three to four-fold higher values were obtained compared to the previously best production strains after purification and activity testing.

After the co-overproduction of GroESL in the AMV-RT production strains had proven to be a positive measure, the other main chaperone system of *E. coli* was additionally co-overproduced in a second step. In addition to the supposed general advantages of this co-overproduction this could compensate for a disadvantage of the GroESL machinery i.e. its exclusion volume of circa 65 kDa (Deuerling E. et al., 1999, Nature 400, 693-696). This should be particularly important for the correct folding of the β -chain of the AMV-RT (93 kDa) provided it cannot be divided into single domains that

are formed independently of one another. The genes DnaK, DnaJ and GrpE were combined in an artificial operon corresponding to the physiological combination (Diamant S. and Goloubinoff P., 1998, Biochemistry 37, 9688-9694; Pierpaoli E.V. et al., 1998, J. Biol. Chem. 273, 6643-6649), whereas the gene for ClpB forms its own transcription unit. Both transcription units were placed under the control of IPTG-inducible T5 promoters in order to coordinate the expression with the genes for the subunits of the AMV-RT.

For technical reasons the cloning process required a number of intermediate steps on the path to the final construct pCHAP-5. Thus the pKKT5 derivatives pCHAP-1 and pCHAP-2 were firstly constructed. pCHAP-1 contains the genetic information for the *dnaKJ* operon from *E. coli* starting with the start codon for *dnaK* up to the stop codon for *dnaJ*; pCHAP-2 carries the artificial transcription unit from the coding regions of the genes for GrpE and ClpB as an insert; the corresponding DNA fragments were amplified by PCR from the genomic DNA of *E. coli* K12KE392. The sequence of the *dnaKJ* operon is shown in SEQ ID NO:14, the corresponding primers used to isolate the *dnaKJ* operon are shown in SEQ ID NO:15 and 16. The sequence of the *grpE* gene is shown in SEQ ID NO:17, the corresponding primers for the isolation of the *grpE* gene are shown in SEQ ID NO:18 and 19. The sequence of the *clpB* gene is shown in SEQ ID NO:20, the corresponding primers for the isolation of the *clpB* gene are shown in SEQ ID NO:21 and 22. In order to construct pCHAP-1 the PCR fragment containing the *dnaKJ* operon was recleaved with *SmaI* and *BamHI* and, as described above, ligated into a vector fragment of pKKT5 which had also been linearized with *SmaI* and *BamHI*. pCHAP-2 was constructed by means of a three-fold ligation with the

EcoRI-PstI fragment of the *grpE* gene, the *PstI-HindIII* fragment of the *clpB*-gene and a vector fragment of pKKT5 linearized with EcoRI and HindIII. p-CHAP-3 in which the *clpB* gene is present alone as a transcription unit, is derived from pCHAP-2 by ligating the *PstI-HindIII* fragment from pCHAP-2 into the vector fragment of pKKT5 linearized with EcoRI and HindIII as described above. Before the ligation reaction the non-compatible ends of the two fragments were prepared with Klenow polymerase (Roche Diagnostics) according to the manufacturer's instructions. pCHAP-4 is a pCHAP-1 derivative whose insert was extended by the *grpE* gene from pCHAP-2 and thus the artificial transcription unit comprises the genes for DnaK, DnaJ and GrpE. As a result of the Shine Dalgarno sequence which is suboptimal in this case, the expression of *grpE* should be reduced compared to pCHAP-2 and thus be better adapted to the expression of *dnaKJ* (Diamant & Goloubinoff, 1998; Pierpaoli et al., 1998). In order to construct pCHAP-4 the EcoRI-AvaI fragment from pCHAP-2 was inserted into the BamHI cleavage site of pCHAP-1 after the non-compatible ends of the two fragments had been prepared with Klenow polymerase (Roche Diagnostics) according to the manufacturer's instructions. The final construct p-CHAP-5 is a pCHAP-4 derivative which contains the insert of pCHAP-3 as additional genetic information. For this the BspLU11I-NdeI fragment in pCHAP-4 was replaced by the BspLU11I-SspI fragment from pCHAP-3 by restriction and ligation as already described several times. In order to ensure the compatibility of the ends, the overhanging ends generated by NdeI were previously filled in with Klenow polymerase (Roche Diagnostics) according to the manufacturer's instructions.

The effect of combining the expression plasmid pAMV $\alpha\beta$ -4

with the various helper plasmids pCHAP-1 to -5 on the overproduction of active AMV-RT was examined. At least under the modified standard expression conditions all helper plasmids considerably increased the previous yields of active AMV-RT and as expected the helper plasmid pCHAP-5 gave the best result. This was confirmed by SDS-PAGE as well as by subsequent purification and activity determination.

5. Example:

Analytical methods

5.1. Test for reverse transcriptase activity (test A)

During the purification, the reverse transcriptase activity in the fractions was detected by means of a non-radioactive test system. The "reverse transcriptase assay non-radioactive" (Roche Molecular Biochemicals, cat. No. 1468120) was used for this. The incubation period was shortened to 30 minutes.

5.2. Test for reverse transcriptase activity (test B)

The specific reverse transcriptase activity of the pools was determined by a radioactive test system. Reverse transcriptase activity was determined in a test volume of 100 µl (50 mM Tris/HCl, pH 8.3 (37°C), 40 mM KCl, 6 mM MgCl₂, 0.5 mM dTTP, 0.04 OD₂₆₀ nm poly (A) x dT₁₅, 0.1 µM [³H]-dTTP). AMV-RT (5 µl) was added in suitable dilutions. After incubating for 10 min at 37°C, the reaction was stopped with 10 % TCA solution (500 µl). The radioactively-labelled product that formed was washed on a nitrocellulose filter after precipitation. The incorporation rate of radioactivity was measured in

a scintillation counter and the RT activity of the sample was calculated. One enzyme unit was defined as the amount of AMV-RT which incorporated 1.0 nMol TMP into acid insoluble product in 10 min at 37°C.

5.3. Test for DNA polymerase

The activity of DNA polymerase from *E. coli* was determined by measuring the nick translation. The DNA polymerase was detected by means of a non-radioactive nick translation test. The nick translation was carried out in a test volume of 50 µl (50 mM Tris/HCl, pH 7.5, 10 mM MgCl₂, 0.1 mM DTE, 28.875 µM DIG-dUTP, 1.444 µM bio-16-dUTP, 95.865 µM dTTP, 20 µM dATP, 20 µM dCTP, 20 µM dGTP, 1 µg pBR322, 1 pg DNaseI). After adding the samples (1 µl) the reaction mixture was incubated for 30 min at 37°C. Afterwards the reaction mixture was transferred to streptavidin-coated microtitre plates. Subsequent treatment and evaluation of the test was carried out analogously to the "reverse transcriptase assay, non-radioactive" (Roche Molecular Biochemicals, Cat. No. 1468120).

5.4 Test for contaminating activities

The test for the presence of contaminating foreign activities was carried out in a solution composed of 10 mM Tris/HCl, pH 7.5, 10 mM MgCl₂, 1 mM DTE.

Suitable samples of the individual enzyme fractions were incubated with the corresponding nucleic acids. So-called nicking activity was detected by incubation with the plasmid pBR322 (1 µg) for 2-16 hours at 37°C. Unspecific nucleases were detected by incubation with

lambda-DNA/*Eco*RI, *Hind*III (1 µg) for 2-16 hours at 37°C. Unspecific RNases were detected by incubation for 2-4 hours at 37°C with MSII-RNA (5 µg).

For the test for contamination by exonucleases, the samples were incubated for 4 hours at 37°C with 4 µg [³H]-labelled DNA and afterwards the released [³H]-labelled nucleotides were determined.

6. Examples:

Purification and functional test

6.1. AMV-RT from *E. coli* LE392 pAMV- α_{lacIq} -Arg/pAMV- β_{dnay} -His construct

6.1.1. Purification

E. coli cells which overexpress both chains of the AMV-RT (see above) were used as the starting material to purify the recombinant AMV-RT.

The AMV-RT was purified at 4°C. The purification was carried out by chromatographic methods after cell lysis and separation of the nucleic acids. The purification process yields a recombinant AMV-RT which is free of contaminating enzyme activities and in RT-PCR has the same functionality as an AMV-RT purified from native material.

Buffer:

buffer A: 50 mM Tris/HCl, pH 7.9, 0.5 M KCl, 0.02 % Triton X-100, 20 % glycerol,
buffer B: 20 mM Tris/HCl, pH 7.9, 0.25 M KCl, 0.02 %

Triton X-100, 10 % glycerol,
buffer C: 20 mM Tris/HCl, pH 7.9, 0.25 M KCl, 0.02 %
Triton X-100, 10 % glycerol, 1 M imidazole,
buffer D: 50 mM Tris/HCl, pH 8.2, 0.1 mM EDTA, 1 mM DTT,
0.02 % Triton X-100, 10 % glycerol,
buffer E: 20 mM potassium phosphate, pH 7.1, 0.1 mM
EDTA, 1 mM DTT, 0.02 % Triton X-100, 10 % glycerol,
storage buffer: 200 mM potassium phosphate, pH 7.2, 2 mM
DTT, 0.2 % Triton X-100, 50 % glycerol.

Cell lysis:

200 ml buffer A was added to ca. 50 g *E. coli* LE392 cells (pAMV- α_{lacIq} -Arg/pAMV- β_{dnar} -His) which were thawed and suspended. Two tablets of Complete (Roche Molecular Biochemicals, cat. No. 1697498) were added to the suspension. Subsequently the cells were lysed by means of ultrasound (Branson sonicator) while cooling (temperature: < 10°C). The degree of lysis of the cell suspension that was achieved was typically 40-50 %.

Precipitation of nucleic acids:

Afterwards the nucleic acids were removed by means of polymin precipitation. 5 ml of a 10 % polymin P solution was added dropwise. If the precipitation was incomplete, further dropwise addition was carried out. After incubation for 30 min at 4°C, centrifugation was carried out (30 min, 13000 rpm, 4°C).

Chromatographic purifications:

Affinity chromatography on a Ni-chelate column:

The clear centrifugation supernatant was diluted with buffer B (1+1) and absorbed to a nickel-loaded chelating sepharose ff column (2.6 cm x 10 cm, Pharmacia) which

had been equilibrated with buffer B, it was then washed with ca. 500 ml buffer B, afterwards with 200 ml buffer B + 10 mM imidazole and 200 ml buffer B + 20 mM imidazole. The enzyme was eluted with a linear gradient of buffer B + 20 mM imidazole and buffer C in a total volume of 500 ml. The flow rate was 5 ml per minute, the fraction size was 20 ml per fraction. The enzyme eluted between 50 mM and 200 mM imidazole. The pool of active fractions was dialysed against buffer D.

Chromatography on heparin-sepharose:

The dialysed pool was subsequently absorbed to a heparin-sepharose ff column equilibrated with buffer D (1.6 cm x 10 cm, Pharmacia) and washed with ca. 200 ml buffer D, then with ca. 200 ml buffer D + 300 mM KCl. The enzyme was eluted with a linear gradient of buffer D + 300 mM KCl and buffer D + 1 M KCl in a total volume of 200 ml. The flow rate was 2.5 ml per min, the fraction size was 10 ml. The AMV-RT eluted at a KCl concentration of 500 mM to 700 mM.

Chromatography on S-sepharose ff:

The RT-active fractions were pooled and dialysed against buffer E. The dialysate was loaded onto a S-sepharose ff column equilibrated with buffer E (1.6 cm x 10 cm, Pharmacia). After washing with ca. 200 ml buffer E, the enzyme was eluted with a linear gradient of buffer E and buffer E + 1 M KCl in a total volume of 400 ml. The flow rate was 2.5 ml per minute, the fraction size was 10 ml.

Chromatography on hydroxylapatite:

The RT-active fractions were pooled and dialysed against buffer E. The dialysate was loaded onto a HA-ultragel column equilibrated with buffer E (1.6 cm x 10 cm,

Biosepra). After washing with ca. 200 ml buffer E, the enzyme was eluted with a linear gradient of buffer E and buffer E + 0.5 M potassium phosphate in a total volume of 400 ml. The flow rate was 2.5 ml per minute, the fraction size was 10 ml.

The RT-active fractions were pooled and dialysed against storage buffer. Application buffer containing SDS and β -mercaptoethanol was added to the purified protein and the sample was denatured by boiling (5 min, 100°C). Subsequently 20 μ l aliquots were analysed by an analytical SDS gel (4-20 %) (Laemmli UK., 1970, Nature 227: 555-557). The α - and β -subunits of AMV-RT were found in equimolar ratios.

The described method yields a stable AMV-RT with an equimolar distribution of the α - and β -subunits. The enzyme obtained is functional in RT-PCR.

6.1.2. Functional test in RT-PCR

The recombinant AMV reverse transcriptase that was obtained was examined in a functional test. The functional test consisted of a reverse transcription (RT) coupled with a polymerase chain reaction (PCR). For this 5 units of the recombinant AMV reverse transcriptase was used like the enzyme mixture of the Titan TM One Tube PCR System (cat. No. 1888382, Roche Molecular Biochemicals). A 1.8 kb fragment of the human dystrophin gene was amplified. 10 ng human muscle RNA was used as a template. The primers (400 nM) were the Dys primer 2rev (5'-GAG TGA ATA CAG TTT GCC CAT GGA TTG-3') and the Dys primer 8for (5'-AAG AAG TAG AGG ACT GTT ATG AAA GAG AAG-3'). The target was amplified in a RT-

PCR using the following program: 50°C for 30 min, 94°C for 2 min followed by 10 cycles (94°C for 10 sec, 58°C for 30 sec, 68°C for 1 min 10 sec) and 20 cycles (94°C for 10 sec, 58°C for 30 sec, 68°C for 1 min 10 sec; + 10 sec/cycle). Subsequently it was incubated for 7 min at 68°C. The reaction products of the RT-PCR were separated after stopping the reaction on a 1 % agarose gel (fig. 1).

Figure 1 shows the amplification product of the RT-PCR having a size of 1.8 kb which was obtained using native purified AMV-RT (lane 2) and AMV-RT that was obtained by recombinant means (lane 3). Lanes 1 and 4 show a DNA molecular weight marker VI (cat. No. 1062590, Roche Molecular Biochemicals).

6.2 AMV-RT from *E. coli* LE392 pAMV $\alpha\beta$ -4 + pCHAP-5 construct

6.2.1 Purification

E. coli LE392 pAMV $\alpha\beta$ -4 + pCHAP-5 cells which overexpress both chains of the AMV-RT (see above) were used as the starting material to purify the recombinant AMV-RT.

The AMV-RT was purified at 4°C. The purification was carried out by chromatographic methods after cell lysis and separation of the nucleic acids. The purification yields a recombinant AMV-RT which is free of contaminating enzyme activities and in RT-PCR has the same functionality as an AMV-RT purified from native material.

Buffer:

buffer A: 50 mM NaPO₄, pH 7.2, 1 M NaCl, 3 mM 2-mercaptopethanol, 10 % glycerol,

buffer B: 50 mM NaPO₄, pH 5.0, 1 M NaCl, 3 mM 2-mercaptopethanol, 10 % glycerol,

buffer C: 50 mM NaPO₄, pH 6.0, 1 M NaCl, 3 mM 2-mercaptopethanol, 10 % glycerol, 0.2 M imidazole,

buffer D: 50 mM NaPO₄, pH 7.7, 1 M NaCl, 3 mM 2-mercaptopethanol, 10 % glycerol, 0.5 M imidazole

buffer E: 50 mM NaPO₄, pH 6.0, 3 mM 2-mercaptopethanol, 10 % glycerol,

storage buffer: 200 mM potassium phosphate, pH 7.2, 2 mM DTT, 0.2 % Triton X-100, 50 % glycerol.

Cell lysis:

Ca. 50 g *E coli* LE392 pAMV- $\alpha\beta$ -4+pCHAP-5 cells were mixed with 400 ml buffer A, thawed and suspended. Two tablets of Complete (Roche Molecular Biochemicals, cat. No. 1697498) were added to the suspension. Subsequently the cells were lysed by means of ultrasound (Branson sonicator) while cooling (temperature: < 10°C). The degree of lysis of the cell suspension that was achieved was typically 40-50 %.

Precipitation of nucleic acids:

Afterwards the nucleic acids were removed by means of polymin precipitation. 5 ml of a 10 % polymin P solution was added dropwise. If the precipitation was incomplete, further dropwise addition was carried out. After incubation for 30 min at 4°C, centrifugation was carried out (30 min, 13000 rpm, 4°C).

Chromatographic purifications:

Affinity chromatography on a Ni-chelate column:

The clear centrifugation supernatant was absorbed to a nickel-loaded chelating sepharose ff column (2.6 cm x 10 cm, Pharmacia) which had been equilibrated with buffer A, it was then washed with ca. 500 ml buffer A, afterwards with 500 ml buffer B and 500 ml buffer C. The enzyme was eluted with buffer D in a total volume of 500 ml. The flow rate was 5 ml per minute, the fraction size was 20 ml per fraction. The pool of active fractions was dialysed against buffer E.

Chromatography on heparin-sepharose:

The dialysed pool was subsequently absorbed to a heparin-sepharose ff column (1.6 cm x 10 cm, Pharmacia) equilibrated with buffer E + 250 mM NaCl and washed with ca. 200 ml buffer E + 250 mM NaCl. The enzyme was eluted with a linear gradient of buffer E + 250 mM NaCl and buffer E + 1 M NaCl in a total volume of 200 ml. The flow rate was 2.5 ml per min, the fraction size was 10 ml. The AMV-RT eluted at an NaCl concentration of 500 mM to 700 mM.

The RT-active fractions were pooled and dialysed against storage buffer. Application buffer containing SDS and β -mercaptoethanol was added to the purified protein and the sample was denatured by boiling (5 min, 100°C). Subsequently 20 μ l aliquots were analysed by an analytical SDS gel (4-20 %) (Laemmli UK., 1970, Nature 227: 555-557). The α - and β -subunits of AMV-RT were found in equimolar ratios (fig. 2, lane 6).

Figure 2 shows an SDS gel with samples from the AMV-RT purification

- lane 1: molecular weight marker
- lane 2: native AMV
- lane 3: cell lysis
- lane 4: Ni-chelate sepharose, wash with buffer C
- lane 5: Ni-chelate pool
- lane 6: rec. AMV-RT final preparation

The described method yields a stable AMV-RT with an equimolar distribution of the α - and β -subunits. The enzyme obtained is functional in RT-PCR.

6.2.2. Functional test in RT-PCR

The recombinant AMV reverse transcriptase that was obtained was examined in a functional test. The functional test consists of a reverse transcription (RT), followed by a polymerase chain reaction (PCR). 10 units of the recombinant AMV reverse transcriptase was used for this. A 8 kb, 10 kb, 12 kb and a 13.5 kb fragment of the human dystrophin gene was amplified.

1 μ g human muscle RNA was used as a template. The primers (400 nM) were the Dys primer 2 for (5'-CAA TCC ATG GGC AAA CTG TAT TCA CTC-3') and Dys primer 5 rev (5'-CGT CCC GTA TCA TAA ACA TTC AGC AGC-3') for 8 kb, Dys primer 8 for (5'-AAG AAG TAG AGG ACT GTT ATG AAA GAG AA-3') and 5 rev for 10 kb, Dys primer 8 for and Dys primer 9 rev (5'-AGC AGG TAA GCC TGG ATG ACT GAC TAG AAG-3') for 12 kb and Dys primer 8 for and 10 rev (5'-AAT CAA TCA ACC AAC CGA AAT CTC ACT CTG-3') for 13.5 kb. The cDNA synthesis was carried out for 60 min at 42°C.

The cDNA synthesis was carried out according to the instructions in the product information for the AMV reverse transcriptase (cat. No. 1495062, Roche Molecular Biochemicals).

The Expand Long Template PCR System (cat. No. 1681834, Roche Molecular Biochemicals) was used for the PCR. The target was amplified using the following PCR program: 94°C for 2 min, followed by 10 cycles (94°C for 10 sec, 60°C for 30 sec, 68°C for 10 min) and 20 cycles (94°C for 10 sec, 60°C for 30 sec, 68°C for 10 min 10 + 10 sec/cycle). Subsequently it was incubated for 5 min at 68°C. After stopping the reaction, the reaction products of the RT-PCR were separated on 1 % agarose gel (fig. 3). Lanes 3 and 6 show a DNA molecular weight marker X (cat. no. 1498037, Roche Molecular Biochemicals).

Figure 3 shows an agarose gel on which the reaction products of the RT-PCR using recombinant AMV-RT were separated; lane 1: 8 kb amplification product, lane 2: 10 kb amplification product, lane 3: DNA length standard X, lane 4: 12 kb amplification product, lane 5: 13.5 kb amplification product, lane 6: DNA length standard X.

SEQUENCE LISTING

<110> Roche Diagnostics GmbH
<120> Method for producing an active heterodimeric AMV-RT in prokaryotic cells
<130> 5272/00/DE

<140>
<141>

<160> 22

<170> PatentIn Ver. 2.1

<210> 1
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 1
gatgactgga attcatgact gttgcgtac atctggct 38

<210> 2
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 2
gatgactgct gcagttatta tgcaaaaaga gggctcgcc 40

<210> 3
<211> 41
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 3
gatgactgct gcagttatta atacgcttga aaggtggctt g 41

<210> 4
<211> 1716
<212> DNA
<213> Avian Myeloblastosis Virus

<400> 4
actgttgcgc tacatctggc tattccgctc aaatggaagc caaaccacac gcctgtgtgg 60
attgaccagt ggccccttcc tgaaggtaaa ctgttagcgc taacgcaatt agtggaaaaaa 120
gaattacagt taggacatat agaaccttca cttagttgtc ggaacacacc tgtctttgtg 180
atccggaagg cttccgggtc ttatcgotta ttgcgtact tgccgcgtgt taacgctaag 240
cttgttcctt ttggggccgt ccaacagggg gcgcgggttc tctccgcgt cccgcgtggc 300
tggcccccga tggccttaga cctcaaggat tgcttctttt ctattcctct tgccgaacaa 360
gatcgcaag ctttgcatt tacgcccccc tctgtgaata accagggccc cgctcgaaga 420
ttccaatgga aggtcttgcc ccaaggatg acctgttctc ccactatctg tcagttgata 480

gtgggtcaaa tacttgagcc cttgcgactc aagcacccat ctctgcgcattt gttgcattat 540
atggatgatc ttttgcttagc cgccctaagt catgatgggt tggaaagcgcc agggggaggag 600
gttatacgta cattggaaaag agccgggttc accatttcgc ctgataaggt ccagagggag 660
cccggagtac aatatcttgg gtacaagttt ggcagttacgt atgtacgcacc cgtaggccctg 720
gtacgagaac ccaggatagc caccttgtgg gatgttcaga agctgggtgg gtcacttcag 780
tggcttgc cagcgctagg aatcccgctt cgaactgtgg gcccccttta tgagcagttt 840
cgagggtcag atccataacga ggcgagggaa tggaaatctag acatgaaaat ggcctggaga 900
gagatcgtgc agctcagcac cactgtgc tttggaaacgtt gggaccctgc cttgcctctg 960
gaaggagcgg tcgttagatg tgaacagggg gcaatagggg tcctggaca gggactgtcc 1020
acacacccaa ggcctatgtt gtgggttattt tccacccaaac ccaccaaggc gtttactgtt 1080
tggtagaaag tggctcaccct tttgattact aagctacgtt cttcgccgtt gcaaacctt 1140
ggcaaggagg ttgatatacctt cctgtgcctt gcatgcattt gggaggacct tccgctcccg 1200
gaggggatcc ttttagccct taggggtttt gcaggaaaaaa tcaggagtag tgacacgcca 1260
tctatttttgc acattgcgcg tccactgcattt gtttctctt aagtgggggtt taccgaccac 1320
cctgtaccgg gacccactgtt ctttaccgcac gcctcctcaa gcacccataa ggggggtggta 1380
gtctggaggg agggcccaag gtggggagata aaagaaaatag ctgatttggg ggcacgtgtt 1440
caacaactgg aagcacgcgc tggccatgcacttctgc tggccgcac aacgcccact 1500
aatgttagtga ctgactctgc gtttgggtgg aaaaatgttac tcaagatggg gcaggaggg 1560
gtcccgctca cagcggcgcc ttttatttttta gaggatgcgt taagccaaag gtcagccatg 1620
gccgcgcgttc tccacgtgcg gagtatttgc gagggtccag gtttttcac agaaggaaat 1680
gacgtggcag atagccaaagc caccatttcaaa gctgttccctt tgagagggc taaagatctc 1716

<210> 5
<211> 2574
<212> DNA
<213> Avian Myeloblastosis Virus

<400> 5
actgttgc tacatctggc tattccgctc aaatggaaagc caaaccacac gcctgtgtgg 60
attgaccagt ggcccccatttc tgaaggtaaa cttgtacgc taacgcattt agtggaaaaaa 120
gaattacagt taggacatata agaaccttca cttagttgtt ggaacacacc tgcgttttgtt 180
atccggagg cttccgggttc ttatcgctt ttgcattgtt tgcgcgtgtt taacgcattt 240
cttggccctt ttggggccgtt ccaacagggg ggcgcgggttc tctccgcgtt cccgcgtggc 300
tggcccttga tggctctaga cctcaaggat tgcttctttt ctattccctt tgcggaaacaa 360
gatcgcgaag cttttgcattt tacgtcccc tctgtgaata accaggcccc cgctcgaaga 420
ttccaatggg aggtcttgcg ccaagggtt acctgttcttcc caactatctt tcaagatggg 480
gtgggtcaaa tacttgagcc cttgcgactc aagcacccat ctctgcgcattt gttgcattat 540
atggatgatc ttttgcttagc cgccctaagt catgatgggt tggaaagcgcc agggggaggag 600
gttatacgta cattggaaaag agccgggttc accatttcgc ctgataaggt ccagagggag 660
cccggagtac aatatcttgg gtacaagttt ggcagttacgt atgtacgcacc cgtaggccctg 720
gtacgagaac ccaggatagc caccttgtgg gatgttcaga agctgggtgg gtcacttcag 780
tggcttgc cagcgcttagt aatcccgctt cgaactgtgg gcccccttta tgagcagttt 840
cgagggtcag atccataacga ggcgagggaa tggaaatctag acatgaaaat ggcctggaga 900
gagatcgtgc agtgcagcac cactgtgc tttggaaacgtt gggaccctgc cttgcctctg 960
gaaggagcgg tggcttagatg tgaacagggg gcaatagggg tcctggaca gggactgtcc 1020
acacacccaa ggcctatgtt gtgggttattt tccacccaaac ccaccaaggc gtttactgtt 1080
tggtagaaag tggctcaccct tttgattact aagctacgtt cttcgccgtt gcaaacctt 1140
ggcaaggagg ttgatatacctt cctgtgc tttgggtgg aaaaatgttac tcaagatggg gcaggaggg 1200
gaggggatcc tggtagccct taggggtttt gcaggaaaaaa tcaggagtag tgacacgcca 1260
tctatttttgc acattgcgcg tccactgcattt gtttctctt aagtgggggtt taccgaccac 1320
cctgtaccgg gacccactgtt ctttaccgcac gcctcctcaa gcacccataa ggggggtggta 1380
gtctggaggg agggcccaag gtggggagata aaagaaaatag ctgatttggg ggcacgtgtt 1440
caacaactgg aagcacgcgc tggccatgcacttctgc tggccgcac aacgcccact 1500
aatgttagtga ctgactctgc gtttgggtgg aaaaatgttac tcaagatggg gcaggaggg 1560
gtcccgctca cagcggcgcc ttttattttta gaggatgcgt taagccaaag gtcagccatg 1620
gccgcgcgttc tccacgtgcg gagtatttgc gagggtccag gtttttcac agaaggaaat 1680
gacgtggcag atagccaaagc caccatttcaaa gctgttccctt tgagagggc taaagatctc 1740
cataccgctc tccatatacgg accccgcgcg ctatccaaag cgtgttaat tttttttttt 1800
caggctaggg aggttgcgtt gacgtggccgtt gatgttgcgtt gatgttgcgtt gatgttgcgtt 1860
ggggtaaacc ctaggggtttt gggacccttca cagatatggc agacagactt tacacttagag 1920
cctagaatgg ctcccccgttc ctggctgcgtt gttactgtgg ataccgcctc atctgcgata 1980
gtcgtaactc agcatggccg tggctcatacg gttgtgcac aacatcattt ggcacgcgtt 2040
atcgccgttt tggaaagacc aaaggccataaaa aaaaacagata atgggtccctg cttcacgtt 2100

aaatccacgc gagagtggct cgcgagatgg gggatagcac acaccaccgg gattccgggt 2160
aattcccagg gtcaagctat ggttagagcgg gccaaccggc tcctgaaaaga taagatccgt 2220
gtgcgtgcgg agggggatgg ctttatgaaa agaatccccca ccagcaaaca gggggacta 2280
ttagccaagg caatgtatgc ccttaatcac tttagcgctg gtgaaaacac aaaaacaccg 2340
atacaaaaac actggagacc taccgttctt acagaaggac ccccggttaa aatacgaata 2400
gagacagggg agtggaaaaa agatgaaac gtgctggct ggggacgagg ttatgcagct 2460
gtgaaaaca gggacactga taaggttatt tggttaccct ctcgaaaagt taaaccggac 2520
atcgcccaaa agatgagggt gactaagaaa gatgaggcga gccctcttt tgca 2574

<210> 6

<211> 572

<212> PRT

<213> Avian Myeloblastosis Virus

<400> 6

Thr Val Ala Leu His Leu Ala Ile Pro Leu Lys Trp Lys Pro Asn His
1 5 10 15

Thr Pro Val Trp Ile Asp Gln Trp Pro Leu Pro Glu Gly Lys Leu Val
20 25 30

Ala Leu Thr Gln Leu Val Glu Lys Glu Leu Gln Leu Gly His Ile Glu
35 40 45

Pro Ser Leu Ser Cys Trp Asn Thr Pro Val Phe Val Ile Arg Lys Ala
50 55 60

Ser Gly Ser Tyr Arg Leu Leu His Asp Leu Arg Ala Val Asn Ala Lys
65 70 75 80

Leu Val Pro Phe Gly Ala Val Gln Gln Gly Ala Pro Val Leu Ser Ala
85 90 95

Leu Pro Arg Gly Trp Pro Leu Met Val Leu Asp Leu Lys Asp Cys Phe
100 105 110

Phe Ser Ile Pro Leu Ala Glu Gln Asp Arg Glu Ala Phe Ala Phe Thr
115 120 125

Leu Pro Ser Val Asn Asn Gln Ala Pro Ala Arg Arg Phe Gln Trp Lys
130 135 140

Val Leu Pro Gln Gly Met Thr Cys Ser Pro Thr Ile Cys Gln Leu Ile
145 150 155 160

Val Gly Gln Ile Leu Glu Pro Leu Arg Leu Lys His Pro Ser Leu Arg
165 170 175

Met Leu His Tyr Met Asp Asp Leu Leu Ala Ala Ser Ser His Asp
180 185 190

Gly Leu Glu Ala Ala Gly Glu Glu Val Ile Ser Thr Leu Glu Arg Ala
195 200 205

Gly Phe Thr Ile Ser Pro Asp Lys Val Gln Arg Glu Pro Gly Val Gln
210 215 220

Tyr Leu Gly Tyr Lys Leu Gly Ser Thr Tyr Val Ala Pro Val Gly Leu
225 230 235 240

Val Ala Glu Pro Arg Ile Ala Thr Leu Trp Asp Val Gln Lys Leu Val
245 250 255

Gly Ser Leu Gln Trp Leu Arg Pro Ala Leu Gly Ile Pro Pro Arg Leu
260 265 270

Met Gly Pro Phe Tyr Glu Gln Leu Arg Gly Ser Asp Pro Asn Glu Ala
275 280 285

Arg Glu Trp Asn Leu Asp Met Lys Met Ala Trp Arg Glu Ile Val Gln
290 295 300

Leu Ser Thr Thr Ala Ala Leu Glu Arg Trp Asp Pro Ala Leu Pro Leu
305 310 315 320

Glu Gly Ala Val Ala Arg Cys Glu Gln Gly Ala Ile Gly Val Leu Gly
325 330 335

Gln Gly Leu Ser Thr His Pro Arg Pro Cys Leu Trp Leu Phe Ser Thr
340 345 350

Gln Pro Thr Lys Ala Phe Thr Ala Trp Leu Glu Val Leu Thr Leu Leu
355 360 365

Ile Thr Lys Leu Arg Ala Ser Ala Val Arg Thr Phe Gly Lys Glu Val
370 375 380

Asp Ile Leu Leu Leu Pro Ala Cys Phe Arg Glu Asp Leu Pro Leu Pro
385 390 395 400

Glu Gly Ile Leu Leu Ala Leu Arg Gly Phe Ala Gly Lys Ile Arg Ser
405 410 415

Ser Asp Thr Pro Ser Ile Phe Asp Ile Ala Arg Pro Leu His Val Ser
420 425 430

Leu Lys Val Arg Val Thr Asp His Pro Val Pro Gly Pro Thr Val Phe
435 440 445

Thr Asp Ala Ser Ser Ser Thr His Lys Gly Val Val Val Trp Arg Glu
450 455 460

Gly Pro Arg Trp Glu Ile Lys Glu Ile Ala Asp Leu Gly Ala Ser Val
465 470 475 480

Gln Gln Leu Glu Ala Arg Ala Val Ala Met Ala Leu Leu Leu Trp Pro
485 490 495

Thr Thr Pro Thr Asn Val Val Thr Asp Ser Ala Phe Val Ala Lys Met
500 505 510

Leu Leu Lys Met Gly Gln Glu Gly Val Pro Ser Thr Ala Ala Ala Phe
515 520 525

Ile Leu Glu Asp Ala Leu Ser Gln Arg Ser Ala Met Ala Ala Val Leu
530 535 540

His Val Arg Ser His Ser Glu Val Pro Gly Phe Phe Thr Glu Gly Asn
545 550 555 560

Asp Val Ala Asp Ser Gln Ala Thr Phe Gln Ala Tyr
565 570

<210> 7
<211> 858
<212> PRT
<213> Avian Myeloblastosis Virus

<400> 7
Thr Val Ala Leu His Leu Ala Ile Pro Leu Lys Trp Lys Pro Asn His
1 5 10 15
Thr Pro Val Trp Ile Asp Gln Trp Pro Leu Pro Glu Gly Lys Leu Val
20 25 30
Ala Leu Thr Gln Leu Val Glu Lys Glu Leu Gln Leu Gly His Ile Glu
35 40 45
Pro Ser Leu Ser Cys Trp Asn Thr Pro Val Phe Val Ile Arg Lys Ala
50 55 60
Ser Gly Ser Tyr Arg Leu Leu His Asp Leu Arg Ala Val Asn Ala Lys
65 70 75 80
Leu Val Pro Phe Gly Ala Val Gln Gln Gly Ala Pro Val Leu Ser Ala
85 90 95
Leu Pro Arg Gly Trp Pro Leu Met Val Leu Asp Leu Lys Asp Cys Phe
100 105 110
Phe Ser Ile Pro Leu Ala Glu Gln Asp Arg Glu Ala Phe Ala Phe Thr
115 120 125
Leu Pro Ser Val Asn Asn Gln Ala Pro Ala Arg Arg Phe Gln Trp Lys
130 135 140
Val Leu Pro Gln Gly Met Thr Cys Ser Pro Thr Ile Cys Gln Leu Ile
145 150 155 160
Val Gly Gln Ile Leu Glu Pro Leu Arg Leu Lys His Pro Ser Leu Arg
165 170 175
Met Leu His Tyr Met Asp Asp Leu Leu Leu Ala Ala Ser Ser His Asp
180 185 190
Gly Leu Glu Ala Ala Gly Glu Glu Val Ile Ser Thr Leu Glu Arg Ala
195 200 205
Gly Phe Thr Ile Ser Pro Asp Lys Val Gln Arg Glu Pro Gly Val Gln
210 215 220
Tyr Leu Gly Tyr Lys Leu Gly Ser Thr Tyr Val Ala Pro Val Gly Leu
225 230 235 240
Val Ala Glu Pro Arg Ile Ala Thr Leu Trp Asp Val Gln Lys Leu Val
245 250 255
Gly Ser Leu Gln Trp Leu Arg Pro Ala Leu Gly Ile Pro Pro Arg Leu
260 265 270
Met Gly Pro Phe Tyr Glu Gln Leu Arg Gly Ser Asp Pro Asn Glu Ala
275 280 285
Arg Glu Trp Asn Leu Asp Met Lys Met Ala Trp Arg Glu Ile Val Gln
290 295 300

Leu Ser Thr Thr Ala Ala Leu Glu Arg Trp Asp Pro Ala Leu Pro Leu
305 310 315 320

Glu Gly Ala Val Ala Arg Cys Glu Gln Gly Ala Ile Gly Val Leu Gly
325 330 335

Gln Gly Leu Ser Thr His Pro Arg Pro Cys Leu Trp Leu Phe Ser Thr
340 345 350

Gln Pro Thr Lys Ala Phe Thr Ala Trp Leu Glu Val Leu Thr Leu Leu
355 360 365

Ile Thr Lys Leu Arg Ala Ser Ala Val Arg Thr Phe Gly Lys Glu Val
370 375 380

Asp Ile Leu Leu Leu Pro Ala Cys Phe Arg Glu Asp Leu Pro Leu Pro
385 390 395 400

Glu Gly Ile Leu Leu Ala Leu Arg Gly Phe Ala Gly Lys Ile Arg Ser
405 410 415

Ser Asp Thr Pro Ser Ile Phe Asp Ile Ala Arg Pro Leu His Val Ser
420 425 430

Leu Lys Val Arg Val Thr Asp His Pro Val Pro Gly Pro Thr Val Phe
435 440 445

Thr Asp Ala Ser Ser Ser Thr His Lys Gly Val Val Val Trp Arg Glu
450 455 460

Gly Pro Arg Trp Glu Ile Lys Glu Ile Ala Asp Leu Gly Ala Ser Val
465 470 475 480

Gln Gln Leu Glu Ala Arg Ala Val Ala Met Ala Leu Leu Leu Trp Pro
485 490 495

Thr Thr Pro Thr Asn Val Val Thr Asp Ser Ala Phe Val Ala Lys Met
500 505 510

Leu Leu Lys Met Gly Gln Glu Gly Val Pro Ser Thr Ala Ala Ala Phe
515 520 525

Ile Leu Glu Asp Ala Leu Ser Gln Arg Ser Ala Met Ala Ala Val Leu
530 535 540

His Val Arg Ser His Ser Glu Val Pro Gly Phe Phe Thr Glu Gly Asn
545 550 555 560

Asp Val Ala Asp Ser Gln Ala Thr Phe Gln Ala Tyr Pro Leu Arg Glu
565 570 575

Ala Lys Asp Leu His Thr Ala Leu His Ile Gly Pro Arg Ala Leu Ser
580 585 590

Lys Ala Cys Asn Ile Ser Met Gln Gln Ala Arg Glu Val Val Gln Thr
595 600 605

Cys Pro His Cys Asn Ser Ala Pro Ala Leu Glu Ala Gly Val Asn Pro
610 615 620

Arg Gly Leu Gly Pro Leu Gln Ile Trp Gln Thr Asp Phe Thr Leu Glu
625 630 635 640

Pro Arg Met Ala Pro Arg Ser Trp Leu Ala Val Thr Val Asp Thr Ala
645 650 655

Ser Ser Ala Ile Val Val Thr Gln His Gly Arg Val Thr Ser Val Ala
660 665 670

Ala Gln His His Trp Ala Thr Ala Ile Ala Val Leu Gly Arg Pro Lys
675 680 685

Ala Ile Lys Thr Asp Asn Gly Ser Cys Phe Thr Ser Lys Ser Thr Arg
690 695 700

Glu Trp Leu Ala Arg Trp Gly Ile Ala His Thr Thr Gly Ile Pro Gly
705 710 715 720

Asn Ser Gln Gly Gln Ala Met Val Glu Arg Ala Asn Arg Leu Leu Lys
725 730 735

Asp Lys Ile Arg Val Leu Ala Glu Gly Asp Gly Phe Met Lys Arg Ile
740 745 750

Pro Thr Ser Lys Gln Gly Glu Leu Leu Ala Lys Ala Met Tyr Ala Leu
755 760 765

Asn His Phe Glu Arg Gly Glu Asn Thr Lys Thr Pro Ile Gln Lys His
770 775 780

Trp Arg Pro Thr Val Leu Thr Glu Gly Pro Pro Val Lys Ile Arg Ile
785 790 795 800

Glu Thr Gly Glu Trp Glu Lys Gly Trp Asn Val Leu Val Trp Gly Arg
805 810 815

Gly Tyr Ala Ala Val Lys Asn Arg Asp Thr Asp Lys Val Ile Trp Val
820 825 830

Pro Ser Arg Lys Val Lys Pro Asp Ile Ala Gln Lys Asp Glu Val Thr
835 840 845

Lys Lys Asp Glu Ala Ser Pro Leu Phe Ala
850 855

<210> 8
<211> 62
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 8
gatgactgga attcatgcgt cgccgtcgcc gtcggcgtcg cactgttgcg ctacatctgg 60
ct 62

<210> 9
<211> 65
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 9
gatgactgga attcatgaga ggcagccacc atcaccatca ccatactgtt gcgctacatc 60
tggct 65

<210> 10
<211> 425
<212> DNA
<213> Escherichia coli

<400> 10
ctgttttggc ggatgagaga agatttcag cctgatacag attaaatcag aacgcagaag 60
cggctgtata aaacagaatt tgcctggcg cagtagcgcg gtggtcccac ctgaccccat 120
gccgaactca gaagtgaaac gccgtagcgc cgatggtagt gtggggctc cccatgcgag 180
atagggaaac tgccaggcat caaataaaac gaaaggctca gtcgaaagac tggccttcc 240
gttttatctg ttgttgtcg gtgaacgctc tcctgatgt gacaaatccg cgggagcgg 300
atgttgaacgt tgccaagcaa cggccggag ggtggcgggc aggacgcccg ccataaactg 360
ccaggcatca aattaagcag aaggccatgc tgacggatgg ccttttgcg tttctacaaa 420
ctctt 425

<210> 11
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 11
aaaactgcag agcagtaagc cggtcataaa a 31

<210> 12
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 12
aaaactgcag cgtgtggat gaagtgtatt a 31

<210> 13
<211> 2155
<212> DNA
<213> Escherichia coli

<400> 13
atcagaattt tttttctttt tcccccttga aggggcgaag cctcatcccc atttctctgg 60
tcaccagccg ggaaaccacg taagctccgg cgtcacccat aacagatacg gactttctca 120
aaggagagtt atcaatgaat attcgccat tgcataatcg cgatgtcg aagcgtaaag 180
aagttgaaac taaatctgct ggcggcatcg ttctgaccgg ctctgcagcg gctaaatcca 240
cccgccgcga atgtgtggct gtcggcaatg gccgtatccct tggaaatggc gaagtgaagc 300
cgctggatgt gaaagttggc gacatcgta tttcaacga tggctacggt gtggaaatctg 360
agaagatcga caatgaagaa gtgttgatca tgcgttgcg aacattctg gcaattgtt 420
aagcgtaatc cgccgcacgac actgaacata cgaatttaag gaataaagat aatggcagct 480
aaagacgtaa aattcggtaa cgacgctcggt gtggaaatgc tgcgcggcgt aaacgtactg 540
gcagatgcag tgaaagttac cctcggtcca aaaggccgta acgtatctt gataaaatct 600
ttcgggtcac cgaccatcac caaagatggt gtttccgtt ctcgttgcg aacactggaa 660
gacaaggatgg aaaaatatggg tgccgcacat gtggaaagaaat ttgcctctaa agcaaaacgac 720
gctgcaggcg acggtaaccac cactgcaccc gtaactggctc aggctatcat cactgaaggt 780
ctgaaagctg ttgtgtggg catgaaatccg atggacatcg aacgtgttgcg cgacaaagcg 840
gttaccgtcg cagttgaaga actgaaagcg ctgtccgtac catgctctga ctctaaagcg 900
attgctcagg ttgttaccat ctccgctaactccgacaaa ccgttagttaa actgatcgt 960

gaagcgatgg acaaagtccgg taaagaaggc gttatcacgg ttgaagacgg taccggctcg 1020
caggacgaac tggacgttgt tgaaggtatg cagttcgacc gtggcttacct gtctccttac 1080
ttcatcaaca agccggaaac tggcgcagta gaactggaaa gcccgttcat cctgctggct 1140
gacaagaaaa tctccaacat ccgcgaaatg ctggcggttc tggaagctgt tgccaaagca 1200
ggcaaacccgc tgctgatcat cgctgaagat gtagaaggcg aagcgctggc aactctggtt 1260
gttaacacca tgcgtggcat cgtgaaagtc gctgcggta aagcacccggg ctgcggcgat 1320
cgtcgtaaag ctatgctgca ggatatcgca accctgactg gcggtaccgt gatctctgaa 1380
gagatcggtt tggagctgga aaaagcaacc ctgaaagacc tgggtcaggae taaacgttgt 1440
gtgatcaaca aagacaccac cactatcate gatggcggtt gtagaaggcg tgaagaaagc 1500
ggccgtgtt ctcagatccg tcagcagatt gaagaagcaa cttctgacta cgaccgtgaa 1560
aaactgcagg aacgcgtac gaaactggca gggggcggtt cagttatcaa agtgggtgct 1620
gctaccgaag ttgaaatgaa agagaaaaaa gcacgcgttg aagatgccct gcacgcgacc 1680
cgtgctcgcc tagaagaagg cgtgggttgt ggtgggttgt ttgcgcgtat ccgcgttagcg 1740
tctaaactgg ctgacctgca tggtcagaac gaagaccaga acgtgggtat caaagttgca 1800
ctgcgtgca tggaaagctcc gctgcgtcag atcgtattga actgcggcga agaaccgtct 1860
gttgggtgtt acaccgtttaa aggccggcgc ggcactactg gttacaacgc agcaaccgaa 1920
gaatacggca acatgatcg catgggtatc ctggatccaa ccaaagtaac tcgttctgct 1980
ctgcagtagc cagttctgt ggctggcctg atgatcacca cgaatgcat gttaccgac 2040
ctgcccggaaa acgatgcagc tgacttaggc gctgctggcg gtatggcg catgggtggc 2100
atggccggca tgatgttaatt gcccctgcacc tcgcagaaat aaacaaccc cccggg 2155

<210> 14
<211> 3139
<212> DNA
<213> Escherichia coli

<400> 14
atgggtaaaa taattggtat cgacctgggt actaccaact ctgtgttagc gattatggat 60
ggcaccactc ctgcgtgtct ggagaacgcc gaaggcgatc gcaccacgcc ttcttatcatt 120
gcctatacc agatgggtga aactctagtt ggtcagccgg ctaaacgtca ggcagttagc 180
aaccgcggaa acactctgtt tgcgattaaa cgcctgattt gtcgcgcctt ccaggacgaa 240
gaagtagacgc gtatgtttc catcatgccg ttcaaaaatta ttgctgtga taacggcgac 300
gcattgggtcg aagttaaagg ccagaaaatg gcacccgcg agatttttc tgaagtgtct 360
aaaaaaatga agaaaaccgc tgaagattac ctgggtgaac cggttaactga agctgttattc 420
accgtaccgg catactttaa cgatgtcag cgtcaggca ccaaagacgc aggccgtatc 480
gctgggtctgg aagttaaacg tatcatcaac gaaccgaccg cagctgcgtt ggcttacgg 540
ctggacaaaag gcactggca cctgtactatc ggggtttatc acctgggtgg tggtaactttc 600
gatatttcta ttatcgaaat cgacgaagtt gacggcgaaa aaaccttcga agttctggca 660
accaacgggtg atacccaccc ggggggtgaa gacttcgaca gccgtctgtat caactatctg 720
gttgaagaat tcaagaaaga tcagggcatt gacctgcgcg acgatccgtt ggcaatgcag 780
cgccctgaaag aagccggcaga aaaagcgaaa atcgaactgtt cttccgtca gcagaccgac 840
gttaacctgc catacatcac tgcagacgcg acgggtccga aacacatgaa catcaaagtg 900
actcgtgcga aactggaaag cctgggtgaa gatctggtaa accgttccat tgagccgtt 960
aaagttgcac tgcaggacgc tggcctgtcc gtatctgtata tcgacgacgt tattctcggt 1020
gggtgggtcaga ctctgtatcc aatgggtcag aagaaatgtt ctgaggtttt tggtaaagag 1080
ccgcgtaaag acgttaaccc ggacgaagct gtaccaatcg gtgtctgtt tcaagggtgg 1140
gttctgactg gtgacgtaaa agacgtactg ctgctggacg ttaccccgct gtctctgggt 1200
atcgaacacca tggccgggtt gatgacgcgc ctgatcgcaaaa aacacaccat tatcccgacc 1260
aagcacagcc aggtgttctc taccgtgaa gacaaccagt ctgcgttaac catccatgtg 1320
ctgcaggggtg aacgtaaacg tgcggctgtat aacaaatctc tgggtcgtt caaccttagat 1380
ggtatcaacc cggcaccgcg cggcatccgc cagatcgaaat ttaccccgat tatcgatgt 1440
gacggtatcc tgcacgtttc cgcgaaagat aaaaacagcg gtaaaagagca gaagatcacc 1500
atcaaggtttt ctctgtgtt gacgtggat gaaatccaga aatgttacg cgacgcgac 1560
gctaacccgcg aagctgaccg taagttgaa gagctgggtac agactcgca ccaggccgac 1620
catctgtgc acacgcaccccg taacgtggat gacgtggat ggcacaaactt gcccggctgac 1680
gacaaaactg ctatcgactc tgcgtgtact gactggaaa ctgctgtat aggtgaagac 1740
aaagccgtt tgcaggacgc aatgtggaa ctggcaccagg tttcccgat actgtatggaa 1800
atcgcccgac agcaacatgc ccacgtggat actgcgggtt ctgatgttcc tgcaaaacaa 1860
gcaaaatggat acgtgttgtt cgacgtggat ttgtggat gacgtggat tcaaaagacaa aaaataatcg 1920
ccctataaaac ggttaattat actgacacgg gcaagggggaa atttcccttc cggccgtgca 1980
ttcatctagg gcaattttaaaa aaaaatggc taagcaagat tattacgaga ttttagggct 2040
ttccaaacaca gccaaggacgt gtgaaatcg aaaggcttac aaacgcctgg ccatgaaata 2100
ccacccggac cgtacccagg gtgacaaaga ggcggaggcc aaatttaaag agatcaaggaa 2160

agcttatgaa gttctgaccg actcgaaaa acgtgcggca tacgatcagt atggccatgc 2220
tgcgttgag caaggtggca tggcgccgg cggtttggc ggccggcgcag acttcagcga 2280
tattttttgtt gacgtttcg gcgatatttt tggcgccggc cgtggtcgtc aacgtgcggc 2340
gcgcggtgc gatttacgtc ataacatgga gctcacccctc gaagaagctg tacgtggcgt 2400
gaccaaagag atccgcattc cgactctgga agagtgtgac gtttgcacg gtagcggtgc 2460
aaaaccaggc acacagccgc agacttgtcc gacctgtcat gtttctggc aggtgcagat 2520
gcgccaggga ttcttcgtc tacagcagac ctgtccacac tgcaggccc gcggtaacgct 2580
gatcaaagat ccgtgcaaca aatgtcatgg tcatggtcgt gttgagcgca gcaaaaacgct 2640
gtccgttaaa atcccgccag gggtgacac tggagaccg atccgtctt cggcgcaagg 2700
tgaagcgccc gaggatggcg caccggcagg cgatctgtac gttcaggttc aggttaaaca 2760
gcacccgatt ttccgagcgtg aaggcaacaa cctgtattgc gaagtccccg tcaacttcgc 2820
tatggcgccg ctgggtggcg aaatcgaagt accgaccctt gatggtcgcg tcaaactgaa 2880
agtgcctggc gaaacccaga cccgtaagct attccgtatg cgcgttaaag gctgtcaagtc 2940
tgcgcgggt ggcgcacagg gtgatttgct gtgcgcgtt gtcgtcggaa caccggtagg 3000
cctgaaccaa aggagcggaaac agctgctgca agagctgca gaaagctcg gtggcccaac 3060
cgccgagcac aacagccgc gctcaaaagag ctctttgtat ggtgtgaaga agtttttga 3120
cgacctgacc cgctaataa 3139

<210> 15
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 15
cccccccggg atgggtaaaa taatttgtat cgac 34

<210> 16
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 16
cgcgggatcc ttatttagcgg gtcaggtcgt caaaaaaaaa 37

<210> 17
<211> 594
<212> DNA
<213> Escherichia coli

<400> 17
atgagtagta aagaacagaa aacgcctgag gggcaagccc cgaaagaaat tatcatggat 60
cagcacgaag agattgaggc agttgagcca gaagcttctg ctgagcagggt ggatccgcgc 120
gatgaaaaag ttgcgaatct cgaagctcgat ctggctgaag cccagacccg tgaacgtgac 180
ggcattttgc gtgtaaaagc cgaaatggaa aacctgcgtc gtcgtactga actggatatt 240
gaaaaagccc acaaattcgc gctggagaaa ttcatcaacg aattgctgcc ggtgattgtat 300
agcctgatc gtgcgcgtgg agtggctgtat aaagctaacc cggatatgtc tgcgatgggt 360
gaaggcattg agctgacgct gaagtgcgtat ctggatgttg tgcgtaaagg tggcggtgaa 420
gtgatcgccg aaactaacgt cccactggac ccgaatgtgc atcaggccat cgcaatgggt 480
gaatctgatc acgttgcgcg aggttaacgtat ctgggcattt tgcagaagg ttataacgctg 540
aatggtcgtat cgttgcgtgc ggcgtatgggtt actgttagcga aagcaaaagc ttaa 594

<210> 18
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 18
cgcggattc atgagtagta aagaacagaa aacg 34

<210> 19
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 19
aaaactgcag ttattaagct tttgctttcg ctacagt 37

<210> 20
<211> 2574
<212> DNA
<213> Escherichia coli

<400> 20
atgcgtctgg atcgtcttac taataaattc cagcttgctc ttgccatgc ccaatcactt 60
gcactcgggc acgacacaacca atttatcgaa ccacttcatt taatgagcgc cctgctgaat 120
caggaagggg gtcgggttag tccttattt acatccgctg gcataaatgc tggccagttg 180
cgcacagata tcaatcaggc attaaatcgt ttaccgcagg ttgaaggtac tgggtggat 240
gtccagccat cacaggatc ggtgcgcgtt ctaatcttt gcgacaagct ggcgcaaaaaa 300
cgtgggtata actttatctc gtcagaactg ttgcgttctgg cggcacttga gtctcgccgc 360
acgctggccg acatcctgaa agcagcaggc ggcaccaccc ccaacattac tcaagcgatt 420
gaacaaatgc gtggaggtga aagcgtgaac gatcaaggtg ctgaagacca acgtcaggct 480
ttgaaaaaat ataccatcga ccttaccgaa cgagccgaac agggcaaact cgatccggtg 540
attggtcgtg atgaaagaaaat tcgcgttacc attcagggtc tgcaacgtcg tactaaaaat 600
aacccggtagt tgattggtag accccggcgtc ggtaaaactg ccatcggtga aggtctggcg 660
cagcgttata tcaacggcga agtgcggaa gggttgaaag gccggccgggt actggcgctg 720
gatatgggcg cgctgggtgc tggggcgaaa tatcgcggtg agtttgaaga acgtttaaaa 780
ggcgtgctta acgatcttgc caaacaggaa ggcaacgtca tccttattt cgacgaatta 840
cataccatgg tcggcgcggg taaagccat ggcgcataatgg acgcggaaa catgtgaaa 900
ccggcgcgtgg cgcgttgta attgcactgc gtaggtgcga cgacgcttga cgaatatcgc 960
cagtacattt aaaaagatgc tgcgtggaa cgtcggttcc agaaagtgtt tggtccgag 1020
ccttctgttg aagataccat tgcgttctg cgtggcctga aagaacgtt cgaattgcac 1080
caccatgtgc aaattactga cccggcaatt gttgcaggg cgacgttgc tcatcgctac 1140
attgctgacc gtcagctgcc ggataaaagcc atgcacgttgc tcgatgaagc agcatccagc 1200
attcgtatgc agattgactc aaaaacaggaa gaactcgacc gactcgatcg tcgttatcatc 1260
cagctaaac tggaaacaaca ggcgttaatg aaagagtctg atgaagccag taaaaaacgt 1320
ctggatatgc tcaacgaaga actgagcgcac aaagaacgtc agtactccga gttagaagaa 1380
gagtgaaag cagagaaggc atgcgtttct ggtacgcaga ccattaaagc ggaactgaa 1440
caggcgaaaaa tcgctatttgc acaggctcgc cgtgtggggg acctggcgcg gatgtctgaa 1500
ctgcaatacg gcaaaatccc ggaactggaa aagcaactgg aagccgcaac gcagctcgaa 1560
ggcaaaacta tgcgtcttttgc ggcgtataaaa gtgaccgcac ccgaaattgc tgaagtgtg 1620
gfcgcgttggc cggggattcc ggtttctcgc atgatggaaa ggcgcgcga aaaactgtg 1680
cgtatggagc aagaactgca ccattcgatcatttgc cgtgtcaga acgaagcggt tgatgcgttgc 1740
tctaaccgttca ttgcgtcgtag ccgtgcgggg ctggccgatc caaatcgccc gattgggttca 1800
ttcctgttcc tcggcccaac tgggtgtggg aaaaacagac tttgttaaggc gctggcgaac 1860
tttatgttttgc atagcgcacca ggcgtatggc cgtatcgatc tgcgttgcgttgc tatggagaaa 1920
cactcggttgcgttgc tggcgttgcgttgc cggggatatgc tgcgttgcgttgc agaagggtggc 1980
tacctgaccg aagcgggtgcg tcgtcgatc tattccgtca tcctgcgttgc tgaagtggaa 2040
aaagcgcaccc cggatgttttgc caacatttgc ttgcgttgcgttgc tggatgtatgg ggcgtctgcact 2100
gacggcgaag ggagaacgtt cgcgttgcgttgc aatacggatc tcattatgc ctctaactc 2160
ggttccgatc tgattcaggc acgcttcgttgc gaaactggatt atgcgcacat gaaagagctg 2220
gtgcgttgcgttgc tggtaaggc taaacttccgttgc ggcgttgcgttgc ttaaccgtat cgatgaagtg 2280
gtggtcttcc atccgcgtggc tgaacagcac attgcctcgttgc tgcgttgcgttgc tcaatgttgc 2340
cgtctgtaca aacgtctggc agaacgttgcgttgc tatggaaatcc acatttgcgttgc gaggcgttgc 2400

aaactgctga gcgagaacgg ttacgatccg gtctatggtg cacgtcctct gaaacgtgca 2460
attcagcagc agatcgaaaa cccgctggca cagcaaatac tgtctggtga attgggtccg 2520
ggtaaagtga ttccgcctgga agttaatgaa gaccggattt tcgcccgtcca gtaa 2574

<210> 21
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 21
aaaactgcag atgcgtctgg atcgtcttac taat 34

<210> 22
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 22
cccgaaaaac ttattactgg acggcgacaa tccggtc 37

Q39E Q42E Q42E Q42E Q42E Q42E